Tempo adjustment of two successive songs

MUS-17

Kevin Machado Duarte

July 5, 2017

Introduction

When playing several songs consecutively, for example during parties, people don't want to have breaks in between the songs and want a smooth transition so that all the songs combined sound like a single song. This task is usually a DJ's job but nowadays systems and algorithms are developed to substitute and eventually outperform the human's capabilities of mixing music.

To create a smooth transition it is not only important to pick an adequate song but also to match the tempi of the two songs, otherwise the transition will be noticeable to the listener. The Japanese corporate KDDI R&D Labs[1] developed an algorithm to solve the problem of adjusting the tempi of two consecutive songs whilst taking the listener's discomfort into consideration.

Problem

The adjustment of the tempi of a song is a procedure that can be reduced to three steps:

1. Tempo estimation/extraction
2. Adjusted tempo computation
3. Tempo modification

In an ideal case, a song has a constant tempo that is predetermined so it can be manually set by someone. However, in most cases one cannot predetermine the tempo throughout the different sections. In that case, it is necessary to first have a tool capable of determining the tempo as accurately as possible. This procedure is known as the tempo estimation or the tempo extraction of a song.

The second step is the problem to which a solution is being presented in this paper.
If no tempo adjustment takes place, the conventional transition consists of cross-fading the volume from the first song to the second song (see Figure 1). This technique involves lowering the volume of the first song while increasing the volume of the succeeding song. This allows a transition from one song to another without having a break in between. This technique is appropriate if the two consecutive songs have the same tempo since no tempo adjustment would be necessary, but causes high discomfort if the difference of the two tempi is great.

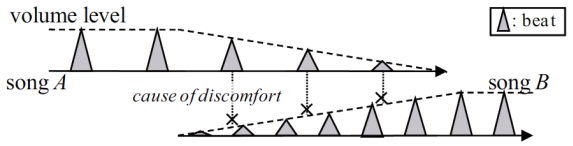

Figure 1: Cross-fade transition from song A to song B [1]

The naive solution to this problem is to adapt one song's tempo to the other tempo, as in Figure 2, and combine it with a cross-fade. Although this method causes less discomfort compared to the simple cross-fade technique, it still causes a lot of discomfort if the two songs have noticeably different tempi.
Real life DJs use a similar technique, however they adjust the second song's tempo instead of the first one, since the first song is already play-

[1] https://www.kddi-research.jp/english

ing and it would be more difficult to adjust it accurately on the fly.

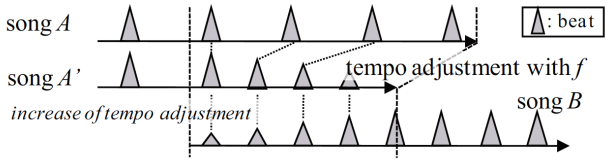

Figure 2: Naive tempo adjustment method [1]

History and Time

Adjusting the tempo of a song is something that has been done manually since the beginning of djing. Djs usually match the second song's tempo to the first one and then proceed with a cross-fade.
However, having this task done by a computer is a subject that has been researched worldwide since the early 2000s. Several solutions exist already, but researchers continue looking for new methods that are able of better adjusting the tempo during a transition between different songs.
The solution illustrated in this paper is mainly based on a paper[1] published by KDDI R&D Labs[2]. Although it dates from 2009, it is one of the most recent papers that covers tempo adjustment for successive songs.

Objective

A lot of research has been done on tempo extraction already and several algorithms exist that are capable of accomplishing this task. A typical extraction method is presented by Gouyon[2] and follows the general scheme presented in Figure 3.

[2]https://www.kddi-research.jp/english

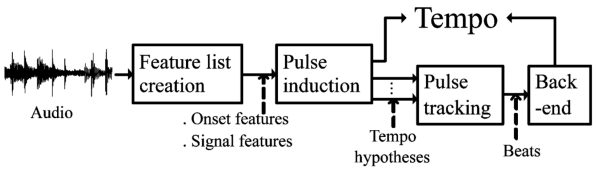

Figure 3: Tempo estimation procedure scheme [2]

Although this method dates back to 2006, it is capable of reaching an accuracy of 80% when estimating the tempo of a song that has little variation throughout it. It is possible to increase the accuracy for example by adding a beat tracking block, however the extraction method will not be analysed in this paper since the focus is situated on the tempo adjustment.

The objective for this paper is to describe a method capable of adjusting the tempi for the transition of two consecutive songs, such that the transition becomes as smooth as possible to avoid a noticeable transition. To do so, it is important for the system to know when a transition is smooth. This is done by evaluating the level of discomfort people experience when a song's tempo is modified and creating a limit to which one can say that as long as the tempo adjustment factor stays in that interval, the transition can be considered as smooth.

Approach

The proposed method [1] consists in creating a smooth song-to-song transition by adjusting the tempi of the two consecutive songs (see Figure 4) instead of trying to match one song's tempo to the other one. This method allows one to keep the tempo of each song closer to the original and once the are combined along with cross-fading, the main volume stays permanently constant. The straight forward way to adjust two tempi would be to simply compute the mean tempo and adjust the song's tempi to match the mean tempo.

Along with the idea of dual tempo adjustment,

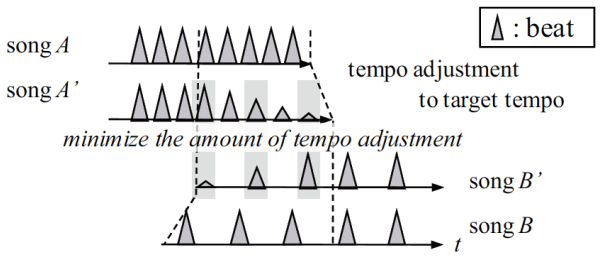

Figure 4: Song-to-song transition with adjustment of the two successive songs [1]

[1] proposes to take the listeners' level of discomfort into consideration. This is accomplished by evaluating the level of discomfort people experience when a song's tempo is modified. During this process, a limit is reached where the tempo adjustment is considered acceptable and produces a smooth transition.

Before creating a formula to define the level of discomfort, an experiment was conducted where a song's tempo would be increased and decreased until people started feeling a discomfort from it. After conducting several tests, it was discovered that listeners weigh the slowdown of a song as a stronger discomfort compared to the speedup.

After this experiment, Eq.1 was set in place to define the level of discomfort of a song's tempo.

$$L_{dc}(f) = \begin{cases} a(f-1), & f > 1 \text{(speed-up)} \\ 0, & f = 1 \text{(no change)} \\ b(\frac{1}{f}-1), & f < 1 \text{(slow-down)} \end{cases} \quad (1)$$

The results of the subject's evaluation in this experiment led to set the weight of the slowdown to cause discomfort to $a = 0.852$ and the weight of the speedup to $b = 1.000$

Once this is set, the algorithm for the actual tempo adjustment can be created.
To determine the target tempo , it will be assumed that two songs A and B have respective tempi T_A and T_B and that $T_A < T_B$.

The first step of the tempo adjustment procedure consists of computing adjusted tempi T'_A for song A as follows:

$$T'_A = 2^C \times T_A \text{ where } C \in \{-2, -1, 0, 1, 2\} \quad (2)$$

Once the T'_As computed, one of the resulted values is picked such that the following equation is fulfilled:

$$C_{opt} = \text{argmin}(|T'_A - T_B|) \quad (3)$$

Lastly, before computing the target tempo and defining the tempi adjustment factors, another parameter needs to be defined:

$$b_{opt} = 2^{C_{opt}} \times T_A \quad (4)$$

Now the target tempo for both songs can be computed by using the previously computed parameters and inputing them into the following formula:

$$T_{tgt} = \frac{(a-b)T_{low} + \sqrt{(a-b)^2 T_{low}^2 + 4abT_{high}T_{low}}}{2a} \quad (5)$$

It is to be noted, that a and b are the weights determined earlier with the level of discomfort. Once the target tempo is set, the only task left before changing the tempo is to define the variation factors as follows:

$$f_{optA} = \frac{T_{tgt}}{b_{opt}}(speed - upfactor) \quad (6)$$

$$f_{optB} = \frac{T_{tgt}}{T_B}(slow - downfactor) \quad (7)$$

After the adjustment factors are computed, the only task left to create the transition is to multiply the original song's tempo with the respective factor. However, to create a smooth speed-up and slow-down effect, the tempo gets progressively modified to avoid an abrupt change of tempo.
Lastly, the adjustment process can be illustrated with Figure 5

It is to be noted that in the case where $T_A > T_B$, one can simply set T_B as the T_A used in the algorithm and analogy switch T_A to T_B.

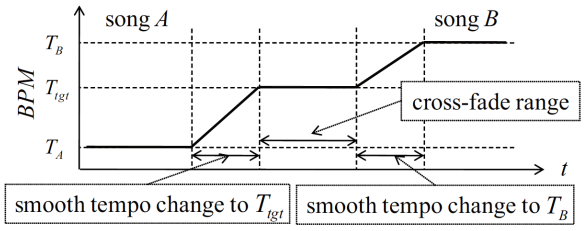

Figure 5: Shifts of tempi of target songs in StS transition [1]

Results

The presented algorithm for tempo adjustment was tested out in a complete automatic dj mixing system and compared to a naive method for tempo adjustment (the first song's tempo is adjusted to the following song).
When looking at the adjustment factors for both methods(see Figure 6), it can be noted that the proposed algorithm keeps the factor closer to 1, which would be equivalent to the original tempo of the song, compared to the naive method where the factor reached a factor of 2(this occurs when the second song's tempo is the double of the first song). Overall, the new method is able to keep the adjustment factor inside the limits of the evaluated level of comfort.

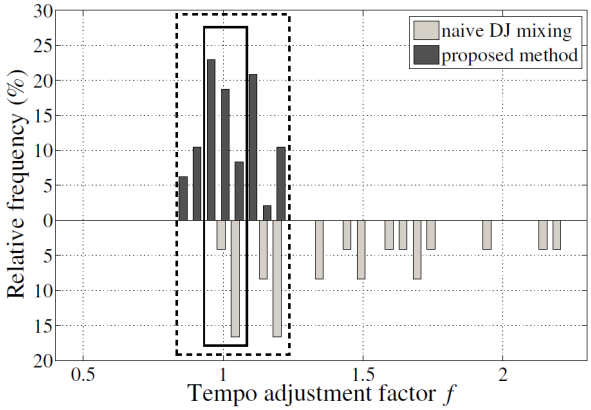

Figure 6: Histogram of the adjustment factors of the proposed and naive DJ mixing method [1]

Along with comparing the technical differences of the two methods, some tests were conducted

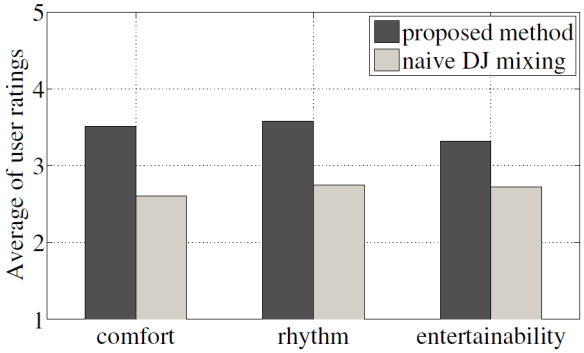

Figure 7: Average of user ratings in proposal and naive DJ mixing [1]

where subjects got to evaluate both methods so that one can compare them together. The test consisted of playing 30 second snippets of songs extracted from *Jamendo*[3] including the chorus and transitioning from song to song with a 5 second intervals for the tempo adjustments section and a cross-fade in between. The subjects then evaluated the comfort, the rhythm and the entertainability of the resulted mixed sounds by rating them on a scale from 1 to 5 with 1 considered as bad and 5 as good. As illustrated in Figure 7, the results of this experiment shows that the new method is able to generate a better comfort, rhythm and entertainability compared to the naive mixing method.

Conclusion

In this paper, a new method for adjusting the tempi of two successive songs was presented. To adapt the proposed method to a new group of listeners, it is best to re-evaluate the weights for the speedup and slowdown that define the level of discomfort as this may vary depending on the listeners and the genres of music they listen to.

Nonetheless, it can be concluded that nowadays it is possible to create a smooth transition from a song to another one just by using a computer and without needing a DJ with experience.

To complete this research, it would be necessary

[3]https://www.jamendo.com/

to compare the presented method with the work of an actual DJ since it would be interesting to know if a system can beat a human in such a task. Although the system might have the technical advantage of computing the ideal tempo adjustment, it is not capable of evaluating it the same way a human does and therefore no conclusion can be taken from this.

References

[1] H. Ishizaki, K. Hoashi, Y. Takishima. *Full-automatic DJ mixing system with optimal tempo adjustment based on measurement function of user discomfort*. ISMIR, 2009, pp. 135-140

[2] F. Gouyon et al. *An Experimental Comparison of Audio Tempo Induction Algorithms*. IEEE ASLP ,2006.

[3] D. Cliff. *Hang the DJ: Automatic Sequencing and Seamless Mixing of Dance-Music Tracks*. 2000,pp. 1-10.

[4] M. Alonso, B. David, G. Richard *Tempo and Beat Estimation of Musical Signals*